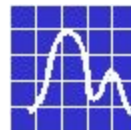
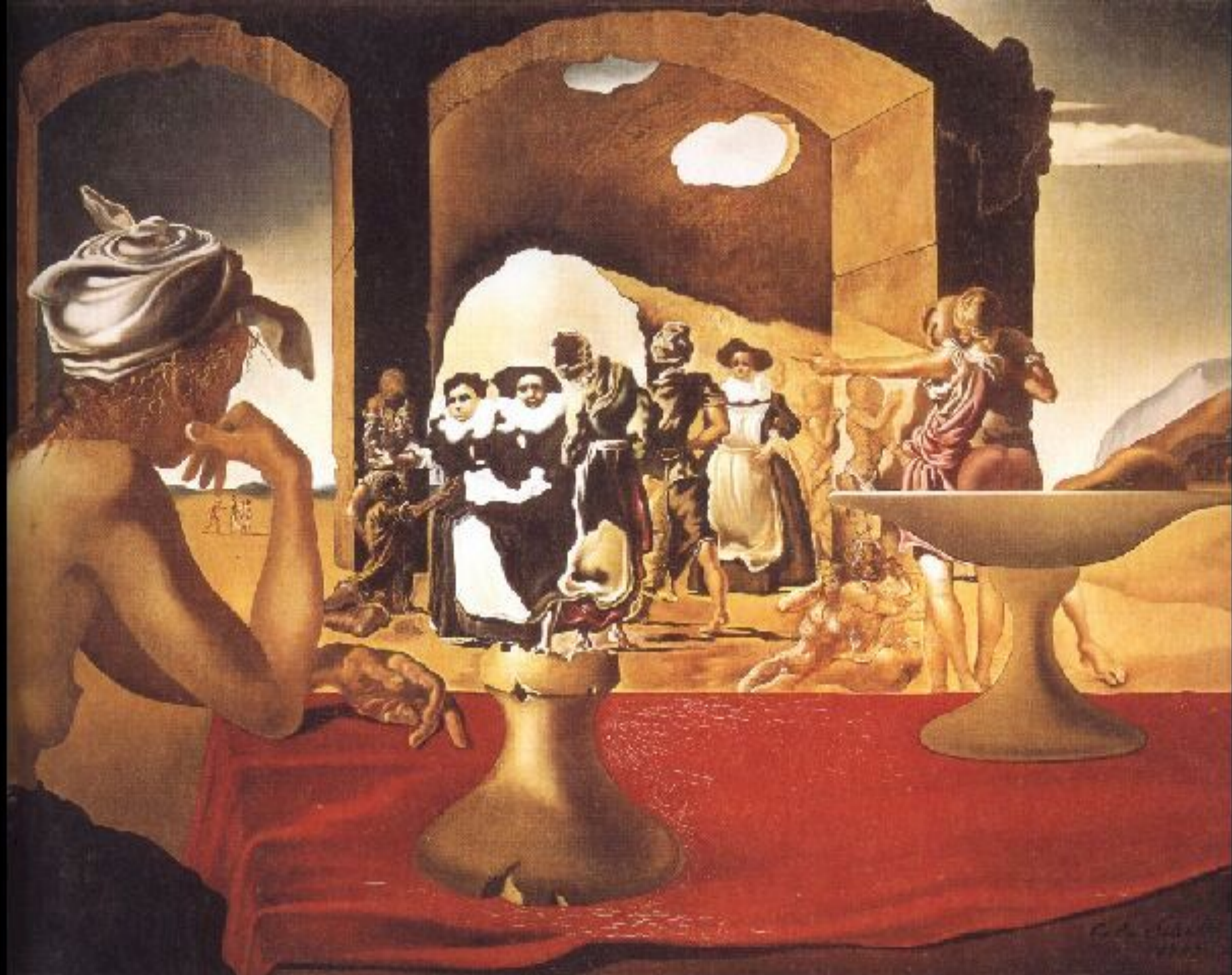


Нейронные сети

Полежаев Илья

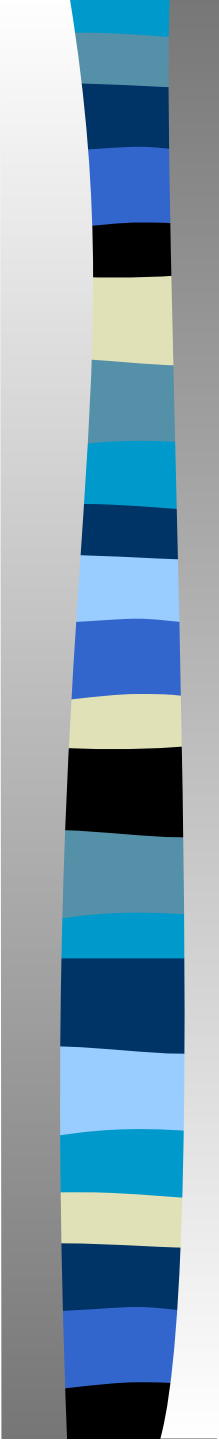


StatSoft® Russia





Физическая аналогия



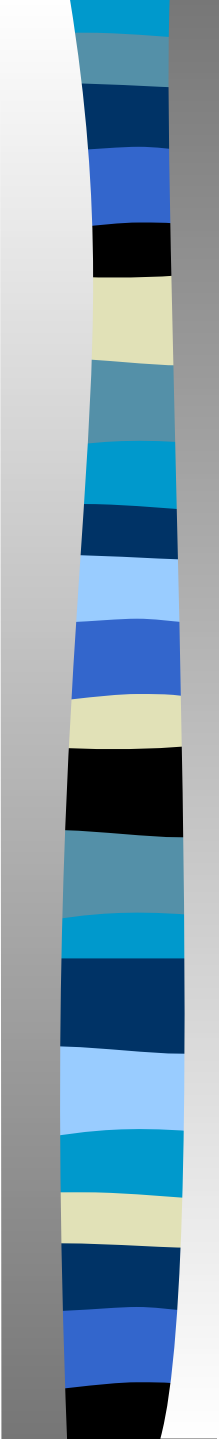
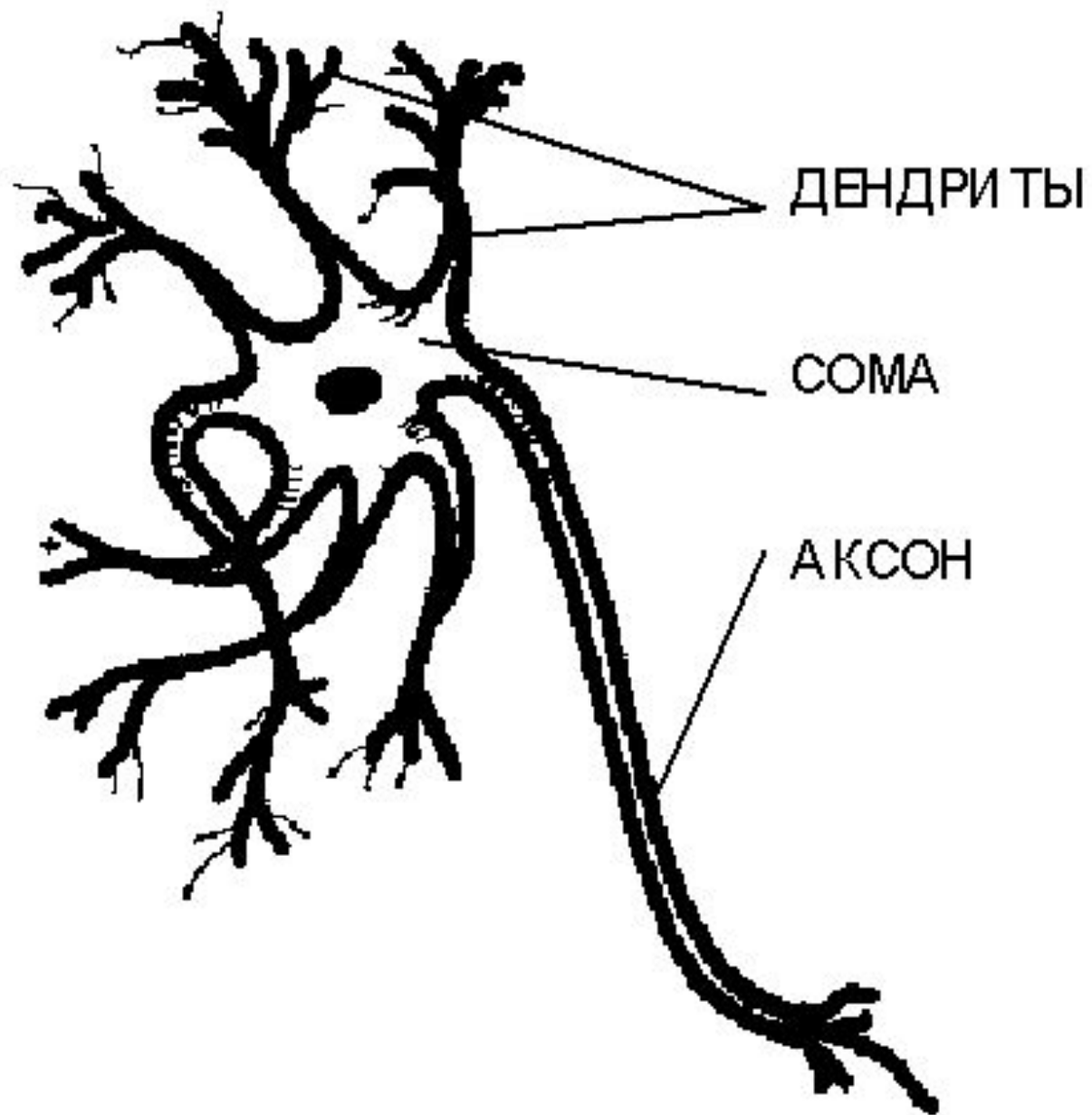
Как он работает?





Современная биология:

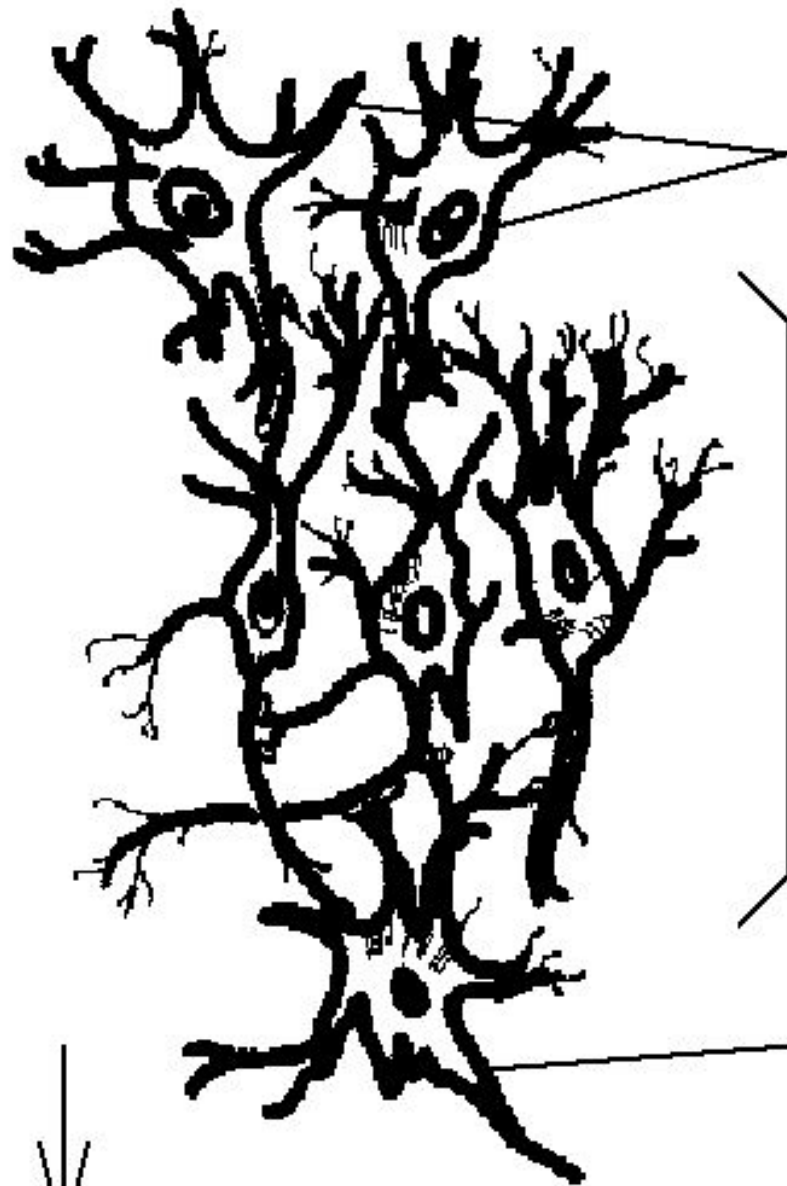
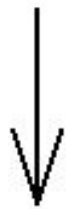
- Клетка - элементарный процессор, способный к простейшей обработке информации
- Нейрон - элемент клеточной структуры мозга
- Нейрон осуществляет прием и передачу информации в виде импульсов нервной активности
- Природа импульсов - электрохимическая



Интересные данные

- Тело клетки имеет размер 3 - 100 микрон
- Гигантский аксон кальмара имеет толщину 1 миллиметр и длину несколько метров
- Потенциал, превышающий 50 мВ изменяет проводимость мембраны аксона
- Общее число нейронов в ЦНС человека порядка 100.000.000.000
- Каждая клетка связана в среднем с 10.000 других нейронов
- Совокупность в объеме 1 мм³ - независимая локальная сеть

ОТ РЕЦЕПТОРОВ



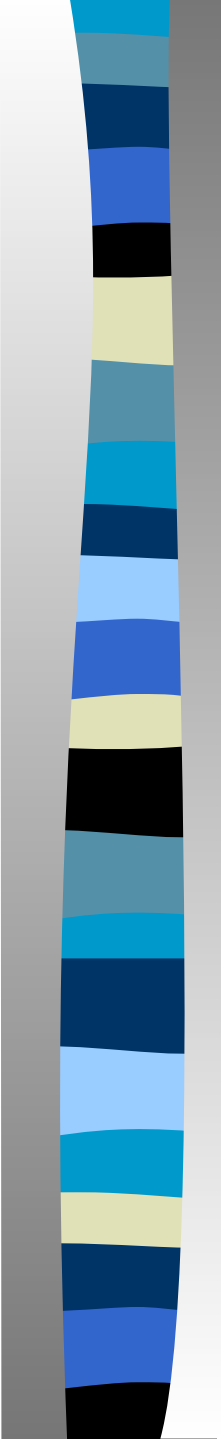
СЕНСОРНЫЕ
НЕЙРОНЫ

НЕЙРОНЫ
ЛОКАЛЬНОЙ
СЕТИ

МОТОРНЫЙ
НЕЙРОН

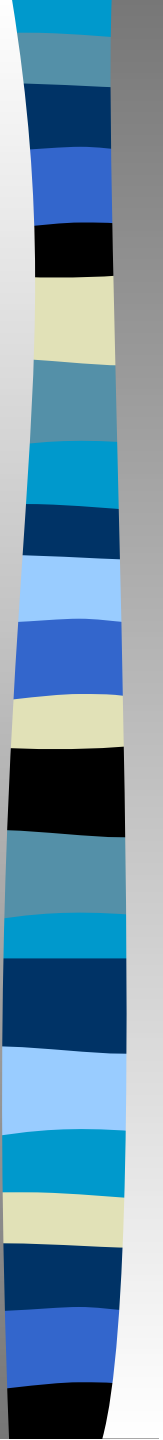


К МЫШЦЕ



Нервная ткань:

- Лишена регенерации
- Её нейроны способны формировать новые отростки и синаптические контакты
- Развитие нейронных ответвлений сопровождается конкуренцией за синаптические участки
- Специфическая изменчивость нейронных сетей лежит в основе их способности к обучению



Формальный нейрон

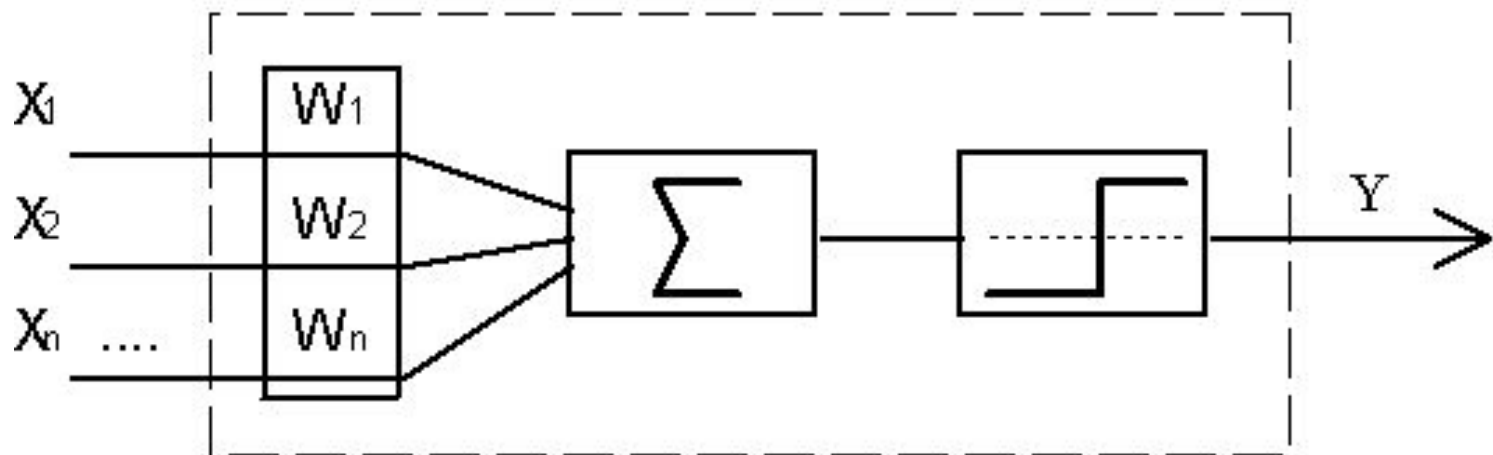
ВХОДНЫЕ
сигналы

Синаптические
веса

Блок
суммирования

Блок нелинейного
преобразования

Выходной
сигнал



$$net = \sum_{i=1}^n W_i X_i$$



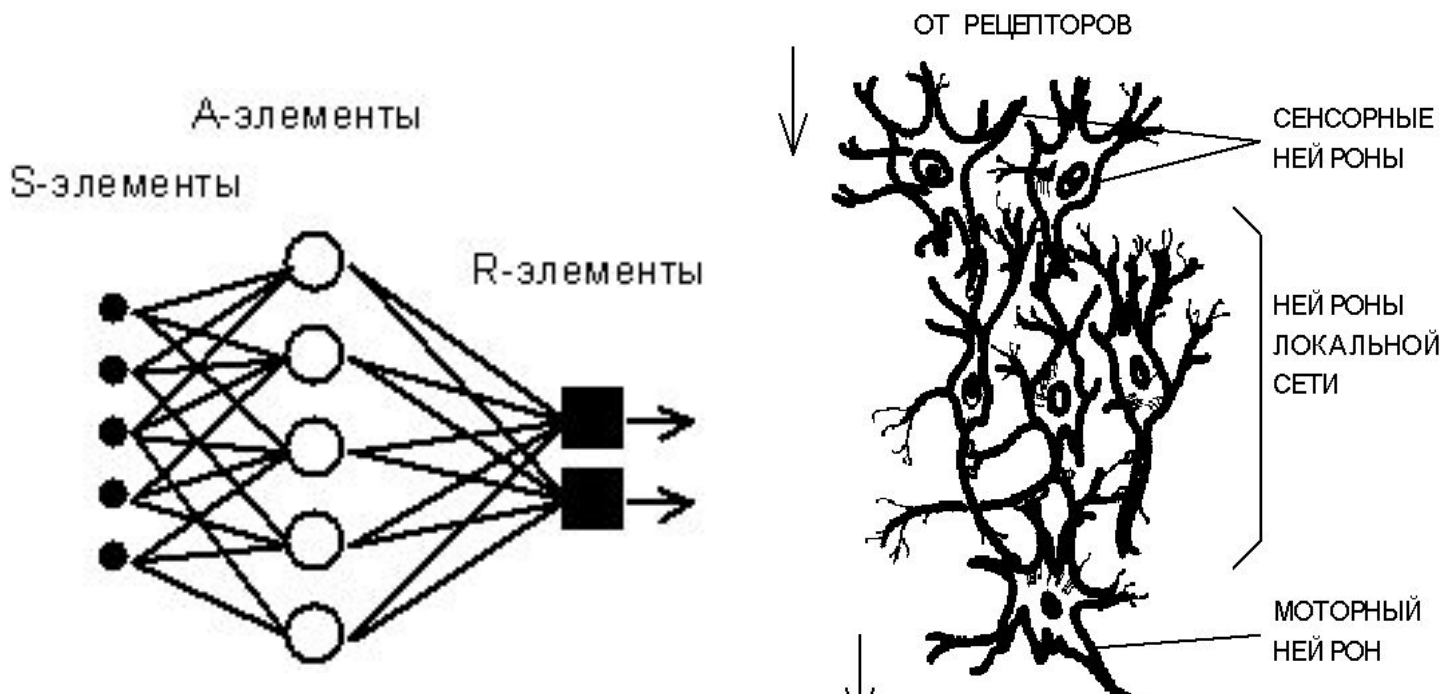
Нелинейное преобразование

Маккалок - Питтс $Y = f(net) = \begin{cases} 1, & net > \Theta \\ 0, & net \leq \Theta \end{cases}$

Линейная $Y = f(net) = \begin{cases} net, & net > \Theta \\ 0, & net \leq \Theta \end{cases}$

Сигмоидальная $Y = f(net) = \frac{1}{1 + \exp(-(net - \Theta))}$

Перцептрон Розенблата



**Розенблат: нейронная сеть рассмотренной архитектуры будет способна к воспроизведению *любой* логической функции.
(неверное предположение)**



Обучение сети

- Обучить нейронную сеть это значит, сообщить ей, чего от нее добиваются.
- Показав ребенку изображение буквы и получив неверный ответ, ему сообщается тот, который хотят получить.
- Ребенок запоминает этот пример с верным ответом и в его памяти происходят изменения в нужном направлении.



Обучение перцептрона

Начальные значения весов всех нейронов полагаются случайными.

$$W(t = 0)$$

Сети предъявляется входной образ x^α , в результате формируется выходной образ.

$$y^{\alpha} \neq y^{\alpha}$$



Обучение перцептрона

Вычисляется вектор ошибки.

делаемой сетью на выходе. $\delta^\alpha = (y^\alpha - \hat{y}^\alpha)$

Идея: изменение вектора весовых коэффициентов в области малых ошибок должно быть пропорционально ошибке на выходе.



Обучение перцептрона

Вектор весов модифицируется по следующей формуле:

$$W(t + \Delta t) = W(t) + \eta X^\alpha \cdot (\delta^\alpha)^T$$

$0 < \eta < 1$ - темп обучения.



Параметры

- Обучение проводится для всех обучающих векторов.
- Один цикл предъявления всей выборки называется **эпохой**.
- Обучение завершается по истечении нескольких эпох, когда вектор весов перестанет значительно меняться.



Возможности применения

Теорема о полноте:

Любая непрерывная функция может быть приближена функциями, вычисляемыми нейронными сетями.

Нейронные сети являются универсальными структурами, позволяющими реализовать любой алгоритм!



Этапы построения сети

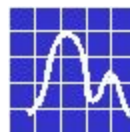
- Выбор архитектуры сети
 - Число входов
 - Функции активации
 - Как соединить нейроны
 - Что взять за вход, что за выход
- Подбор весов (обучение сети)
- Построить вручную
- Воспользоваться пакетом нейросетевого моделирования



STATISTICA

Neural Networks

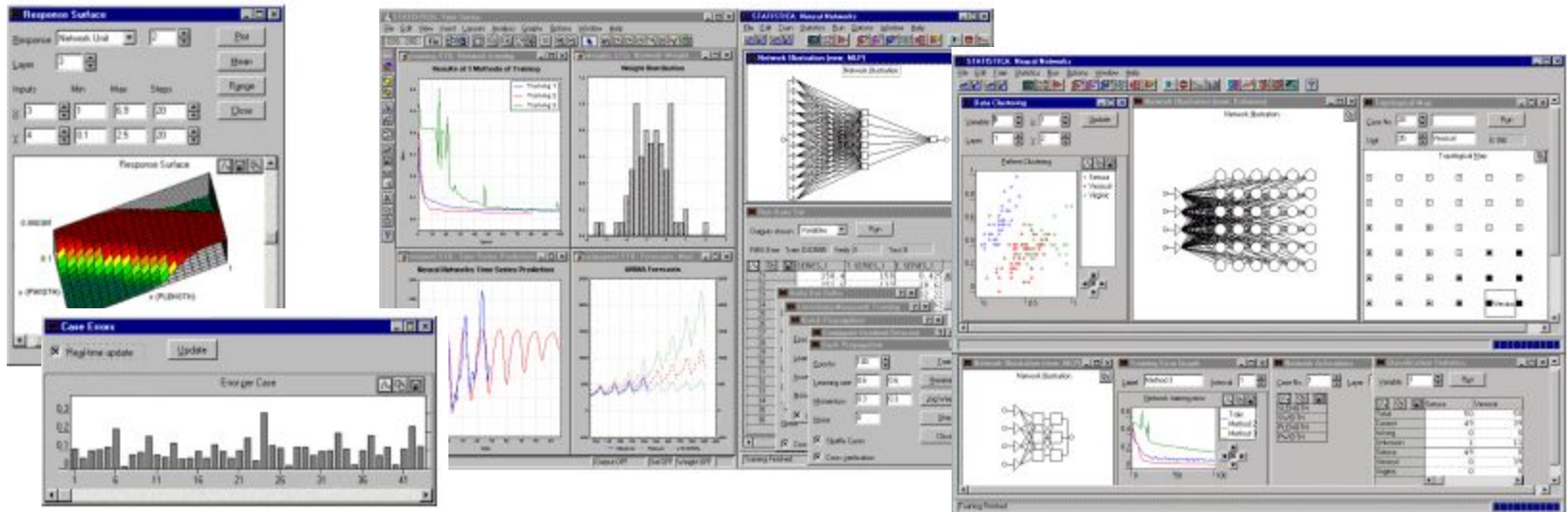
Программный пакет для создания и обучения нейронных сетей и работы с нейросетевыми моделями



StatSoft® Russia

STATISTICA Neural Networks

- Исключительная простота в работе
 - ✓ Советник по конструированию сети
 - ✓ Мастер решения задач
- Богатые средства визуализации





STATISTICA Neural Networks: работа с данными

- Структура таблиц исходных данных:
 - числовые и номинальные переменные;
 - входные и выходные переменные;
 - подмножества наблюдений.
- Импорт файлов различных форматов, использование буфера обмена.
- Подготовка данных: встроенные алгоритмы пре- и пост-процессирования.



STATISTICA Neural Networks: построение сетей

- Создание и сохранение наборов сетей.
- Выбор типа сети:
 - многослойные персептроны (MLP);
 - радиальные базисные функции (RBF);
 - вероятностные и обобщенно-регрессионные сети (PNN и GRNN);
 - сети Кохонена.
- Задание функции ошибок, функций активации и PSP-функций различных слоев.
- Доступ к весам всех нейронов сети.



STATISTICA Neural Networks: обучение сетей

- Большой выбор алгоритмов обучения:
 - обратное распространение ошибки;
 - спуск по сопряженным градиентам;
 - квази-ньютоновский и Левенберга-Маркара;
 - метод псевдообратных матриц.
- Использование кросс-проверки.
- Задание условий остановки.
- Контроль за процессом обучения с помощью графика среднеквадратичной ошибки и гистограммы ошибок наблюдений.



STATISTICA Neural Network: дополнительные функции

- Генетический алгоритм отбора входных данных
- Нелинейное понижение размерности
- Регуляризация весов по Вигенду
- Анализ чувствительности
- Введение матрицы потерь
- Операционные характеристики



STATISTICA Neural Networks: создание приложений

- Взаимодействие с системой *STATISTICA*: передача данных и графиков.
- Встроенный интерфейс прикладного программирования (API) для создания приложений в среде Visual Basic и C++.
- Новая функция - генератор программного кода на языке Си.

Прогнозирование результатов выборов президента США





Условия моделирования

- Предвыборные компании кандидатов отработаны добросовестно
- Все участники сделали все возможное
- Выбор практически предопределяется лишь объективными признаками?
- Прогноз составлялся в 1992 году по данным выборов начиная с 1864



Входные данные

- Правящая партия у власти более 1 срока?
- Правящая партия получила больше 50% на прошлых выборах?
- В год выборов была активна третья партия?
- Была серьезная конкуренция при выдвижении кандидата от правящей партии?



Входные данные

- Кандидат от правящей партии был президентом в год выборов?
- Был ли год выборов временем спада или депрессии?
- Был ли рост среднего национального валового продукта на душу населения более 2,1%?
- Произвел ли правящий президент существенные изменения в политике?



Входные данные

- Во время правления были существенные социальные волнения?
- Администрация правящей партии виновна в серьезной ошибке или скандале?
- Кандидат правящей партии - национальный герой?
- Кандидат оппозиционной партии - национальный герой?



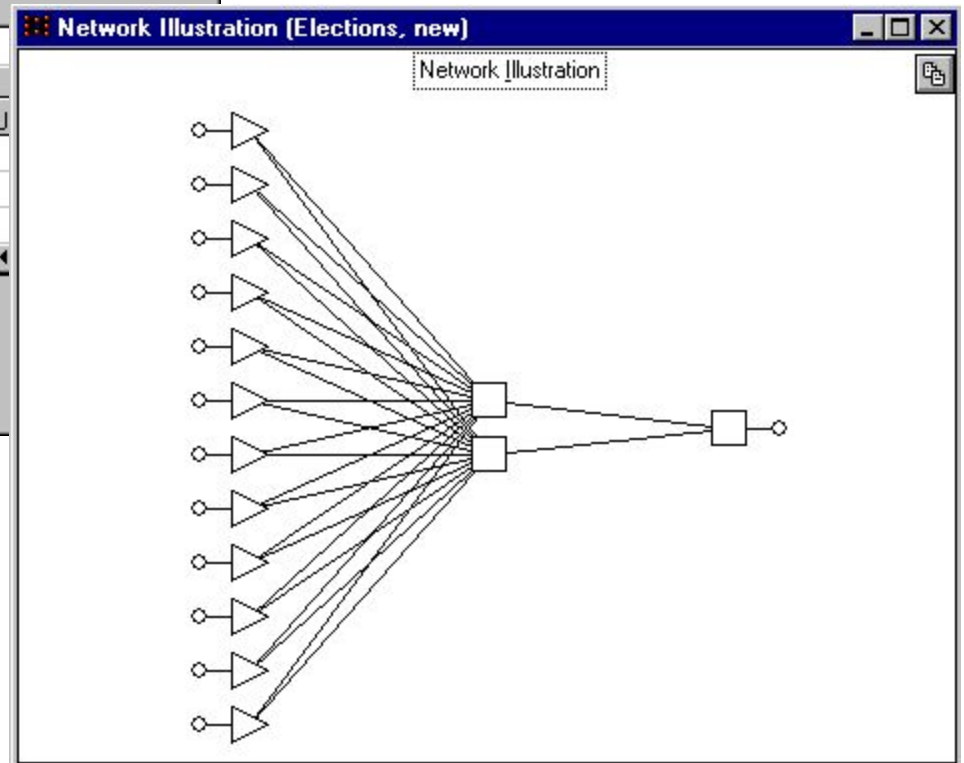
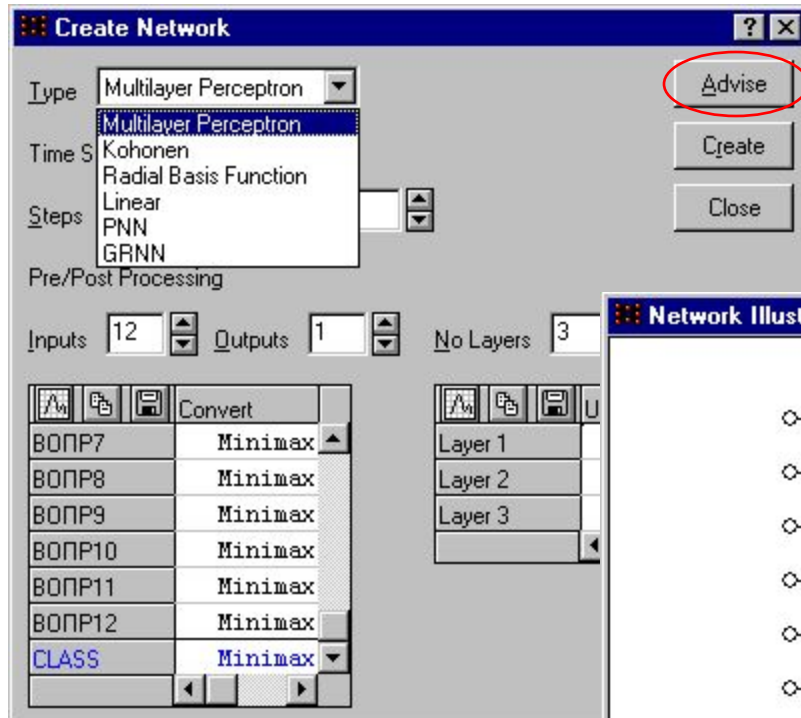
Data Set Editor (Elections)

Variables: 12 | 1 | Cases: 25 | 5 | 1

	ВОПР8	ВОПР9	ВОПР10	ВОПР11	ВОПР12	CLASS
01	1	1	0	0	0	1
02	1	1	0	1	0	1
03	0	0	0	1	0	1
04	1	0	0	0	0	1
05	0	0	0	0	0	1
06	0	0	0	0	1	1
07	0	0	0	1	0	1
08	1	0	0	0	1	1
09	1	0	0	0	0	1
10	1	0	1	0	0	1
11	0	0	0	0	0	1
12	1	0	0	1	0	1
13	1	0	0	1	0	1
14	1	0	0	1	0	1
15	1	0	0	0	0	1
16	0	0	0	1	0	1
17	0	0	0	0	0	1
18	0	1	0	0	0	1
19	0	1	0	0	0	2

0:00:00

Создание сети



Обучение

Back Propagation [?] [X]

Epochs: [Train]

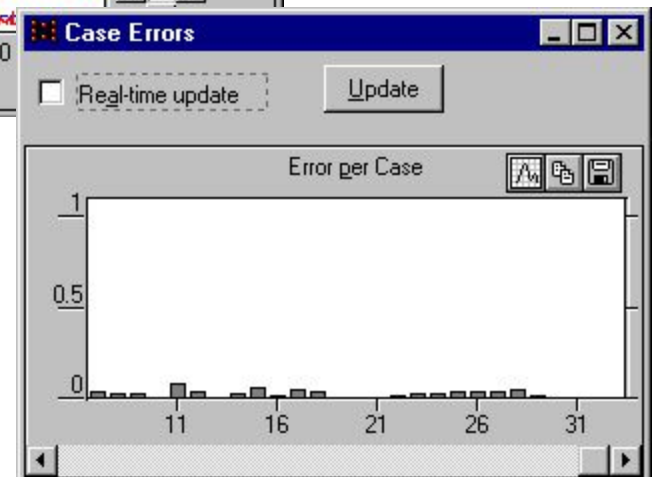
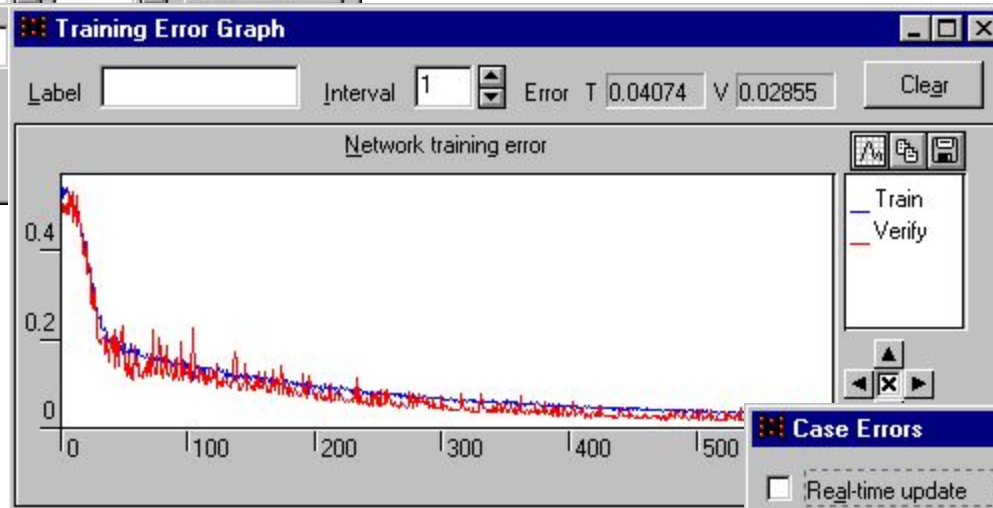
Learning rate: [Reinitialize]

Momentum: [Jog Weights]

Noise:

Shuffle Cases

Cross verification





Data Set Editor (Elections)

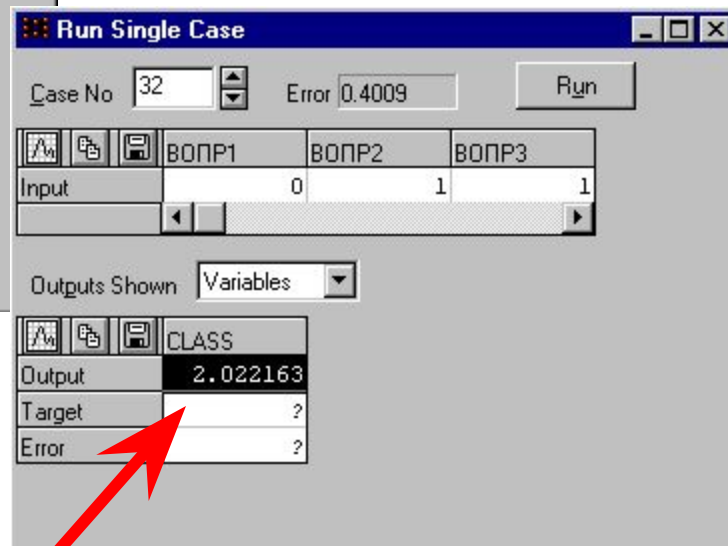
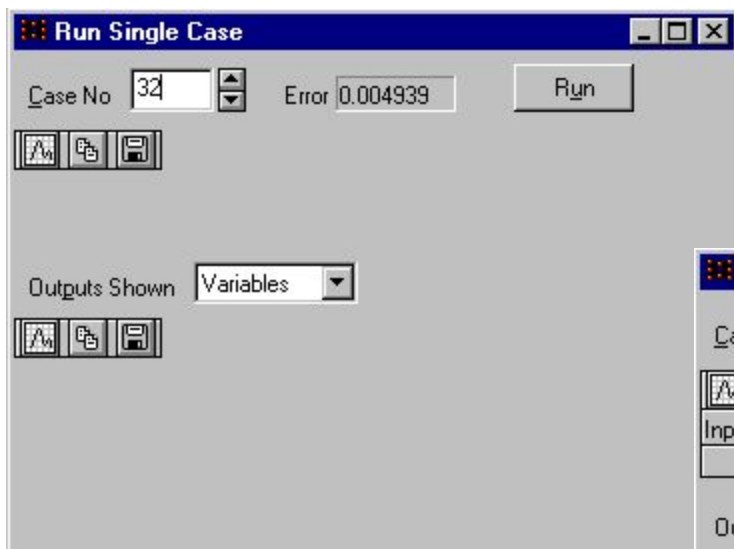
Variables 12 | 1 Cases 27 | 5 | 0

	год	ВОПР1	ВОПР2	ВОПР3	ВОПР4
19	1860	1	0	1	1
20	1872	1	1	0	1
21	1884	1	0	0	1
22	1892	0	0	1	0
23	1896	0	0	0	1
24	1912	1	1	1	1
25	1920	1	0	0	1
26	1932	1	1	0	0
27	1952	1	0	0	1
28	1960	1	1	0	0
29	1968	1	1	1	1
30	1976	1	1	0	1
31	1980	0	0	1	1
32	1992	0	1	1	1
33		1	1	0	0

- Training
- Verification
- Test
- Ignore
- Name/Number
- Cut
- Copy
- Paste
- Clear

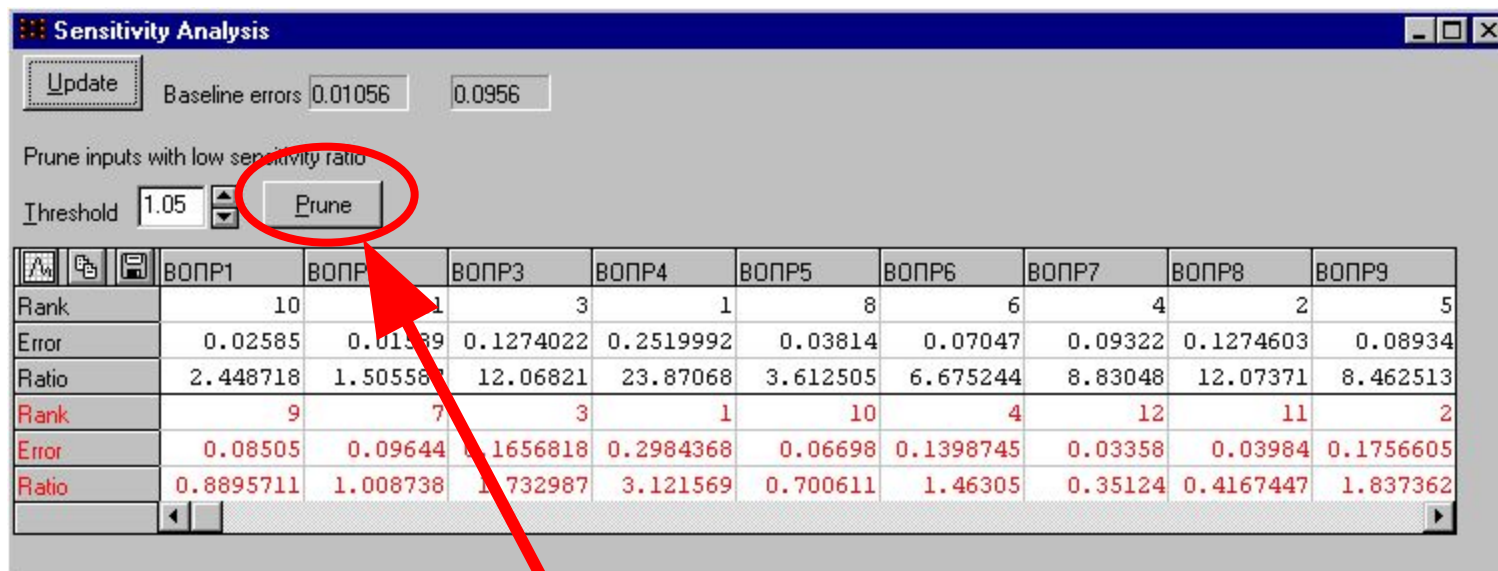
Активизируем случай 1992 года

Прогноз



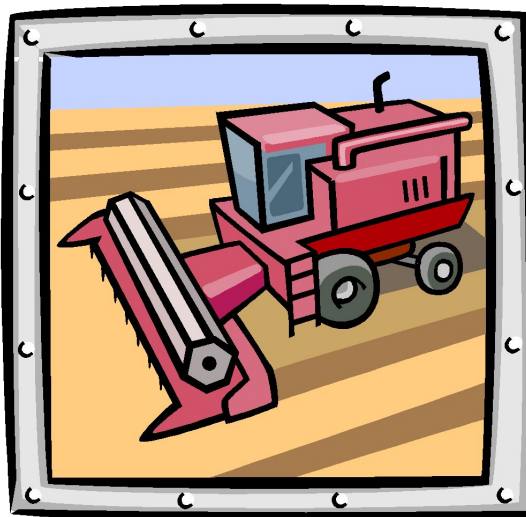
**Результат = 2 -
прогнозируется
победа
кандидата
из оппозиции**

Анализ чувствительности



**Нажатие этой кнопки
автоматически исключает
незначимые переменные
из анализа**

На основе экспертных данных
выявить факторы, наиболее
влияющие на прибыль
предприятия





Представлены факторы

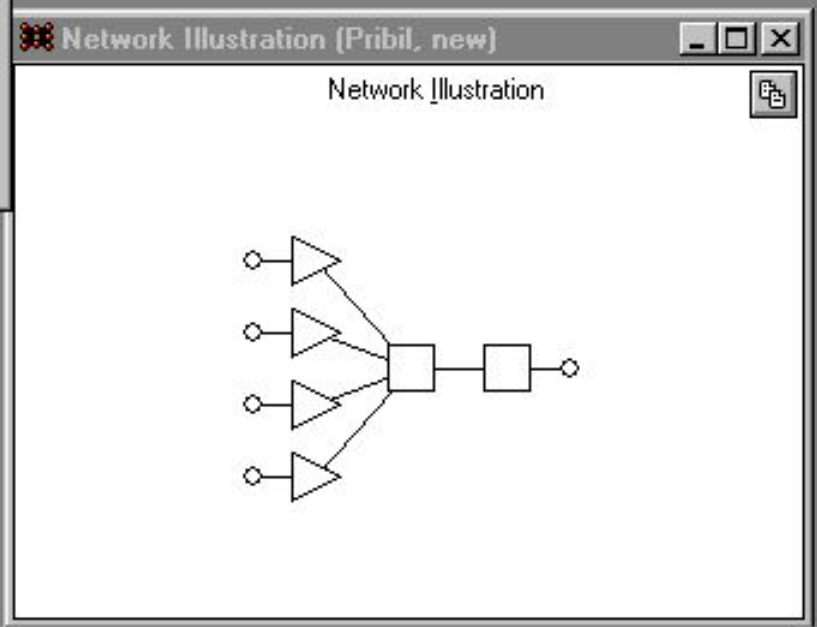
- Затраты на материалы
- Объем зарплаты
- Производительность труда
- Курс доллара США



Data Set Editor (Pribil)

Variables: 4 1 Cases: 12 0 0

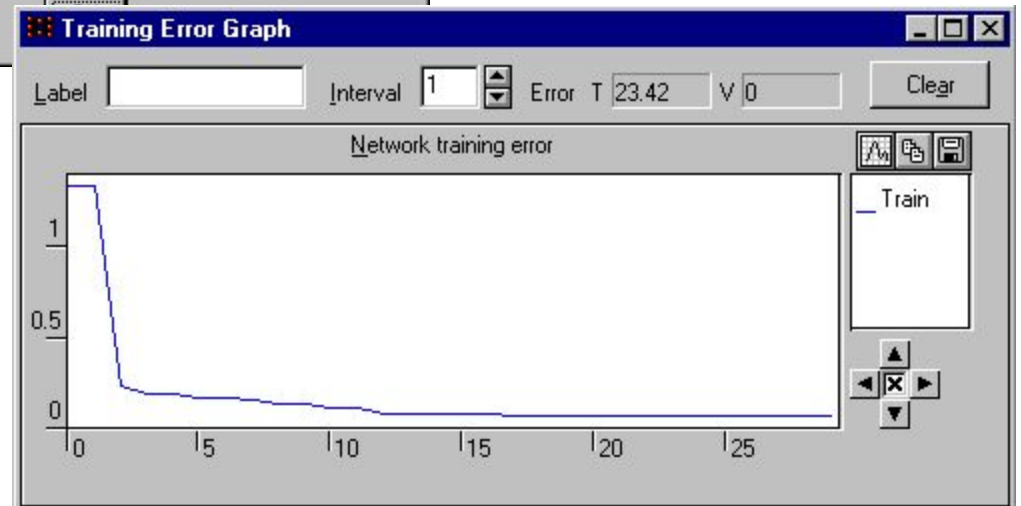
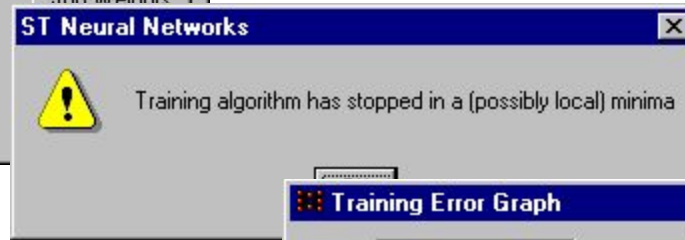
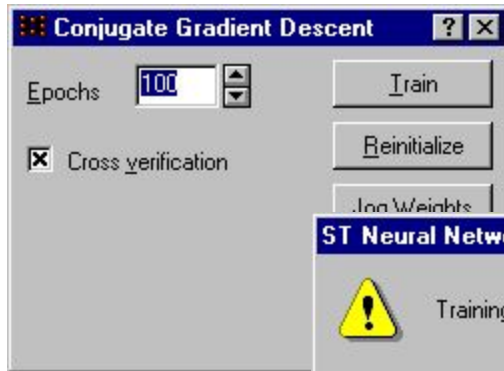
	PAY	PERFOMAN	RATE\$	DIVIDENT
01	122	6003	6	99.2
02	126	4474	6.1	29.1
03	126	4474	6.1	-41.6
04	104	4798	6.1	51.3
05	112	5273	6.2	111.9
06	74	3410	6.2	8.4
07	102	4237	6.2	54.7
08	76	2673	7.9	-3.6
09	123	4406	16.1	-45.4
10	117	4711	16	30.1
11	107	4833	17.9	65.5



Edit the Current Neural Network

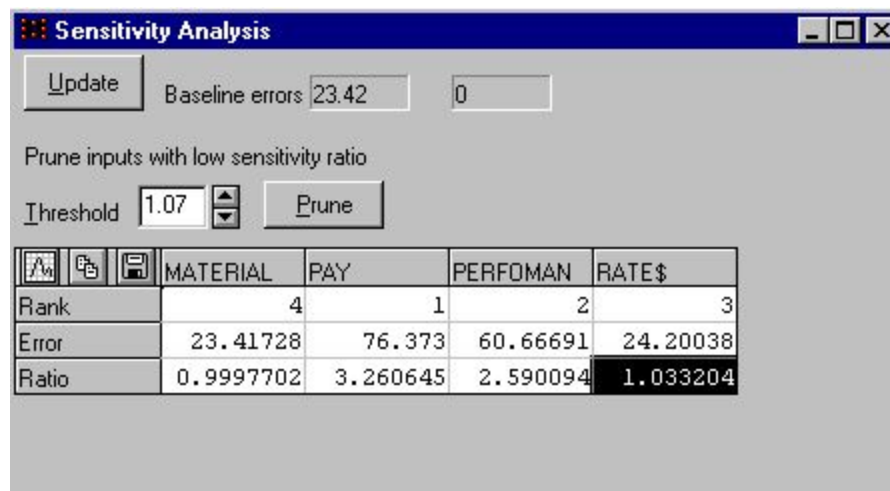
0:00:00

Обучение



**Регрессия
построена**

Анализ чувствительности



	MATERIAL	PAY	PERFORMAN	RATE\$
Rank	4	1	2	3
Error	23.41728	76.373	60.66691	24.20038
Ratio	0.9997702	3.260645	2.590094	1.033204

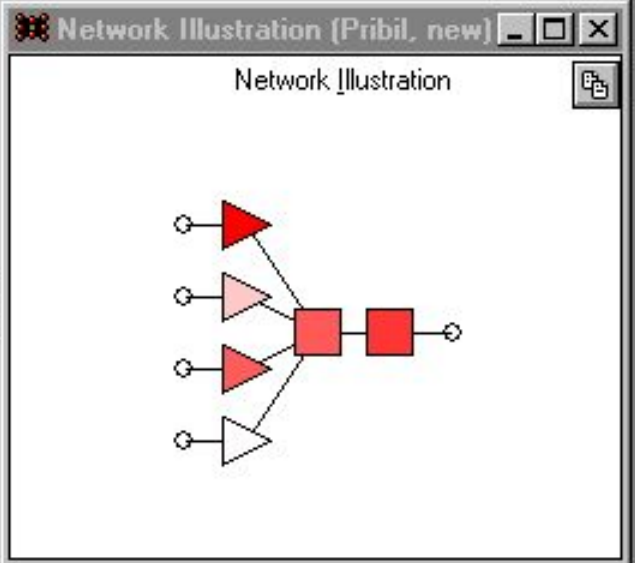
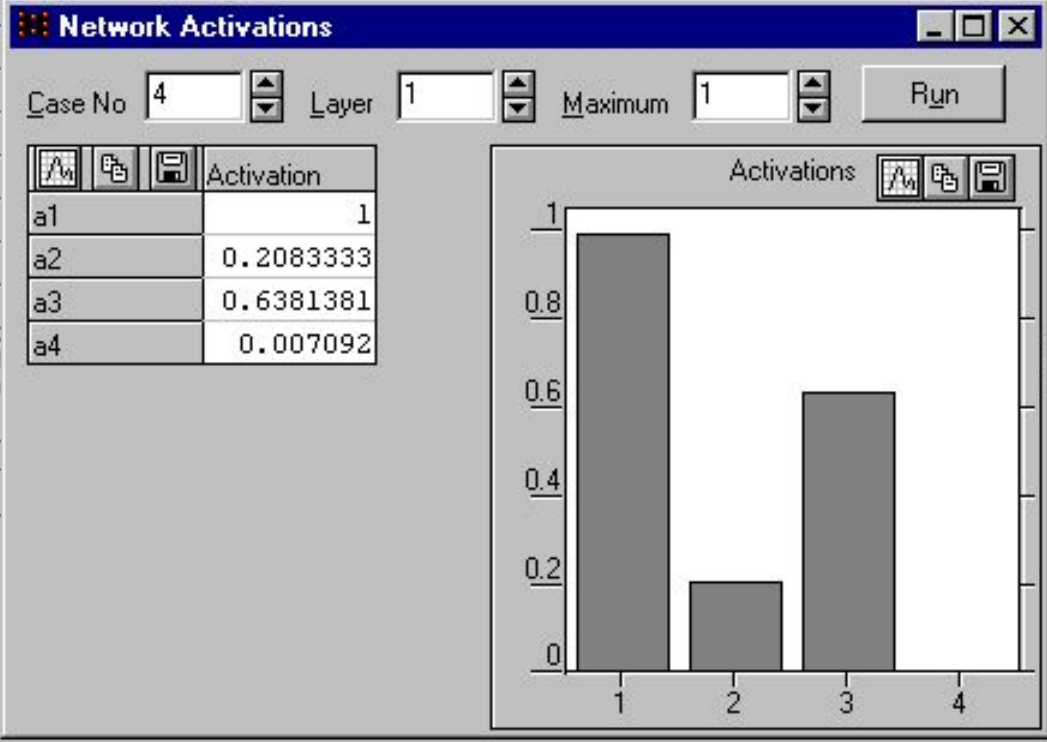
**Объем зарплаты и
производительность
труда сильно влияют
на прибыль предприятия**



Data Set Editor (Pribil)

Variables: 4, 1 Cases: 12, 0, 0

	MATERIAL	ZARPLATA	PROIZVOD	KURS\$	PRIBIL
01	53	122	6003	6	99.2
02	58	126	4474	6.1	29.1
03	44	126	4474	6.1	-41.6
04					
05					
06					
07					
08					
09					
10					
11					



**Анализ и прогнозирование
объема продаж сетей
автозаправочных
станций в США**



График временного ряда





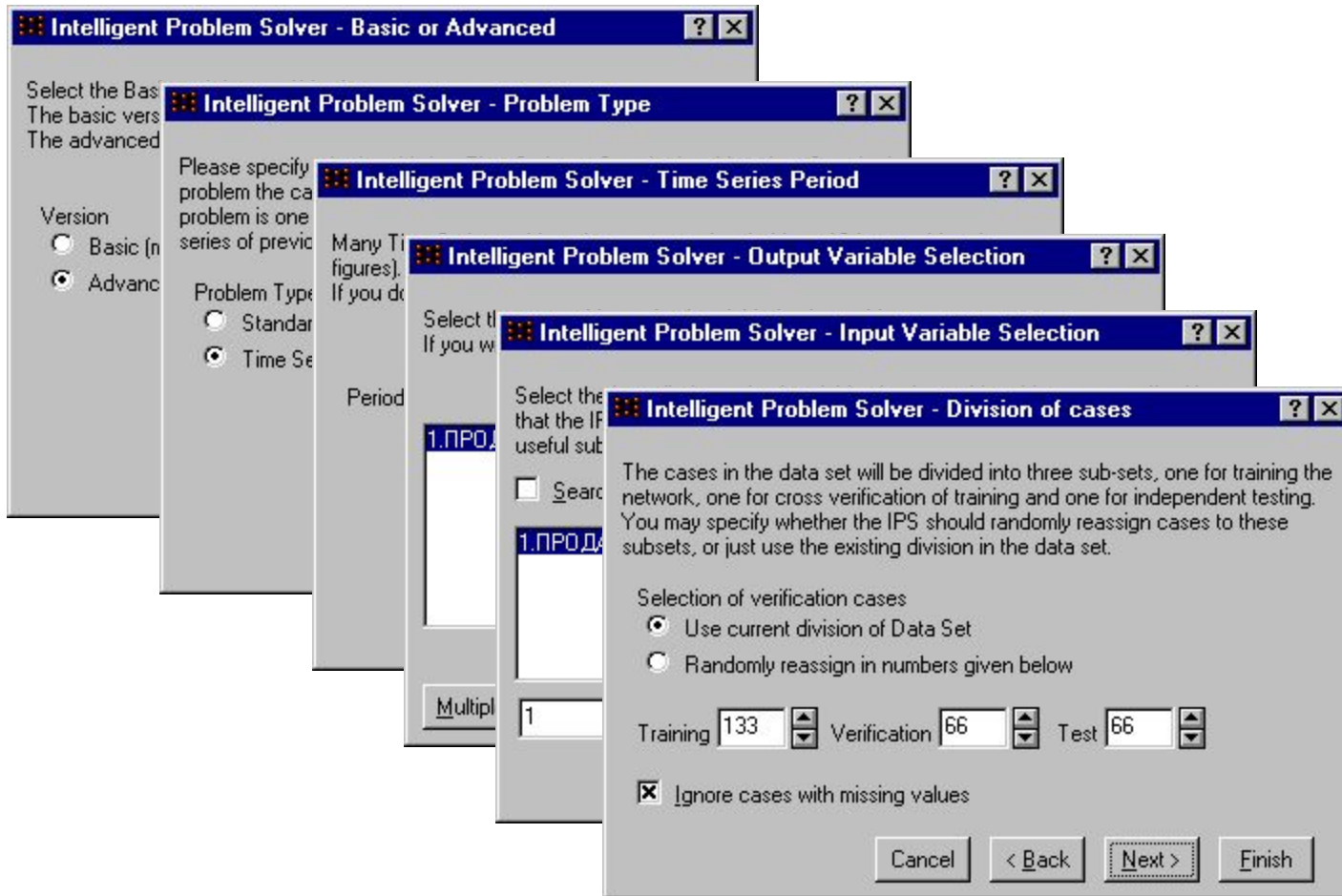
Data Set Editor (Retail)

Variables 1 1 Cases 133 120 0

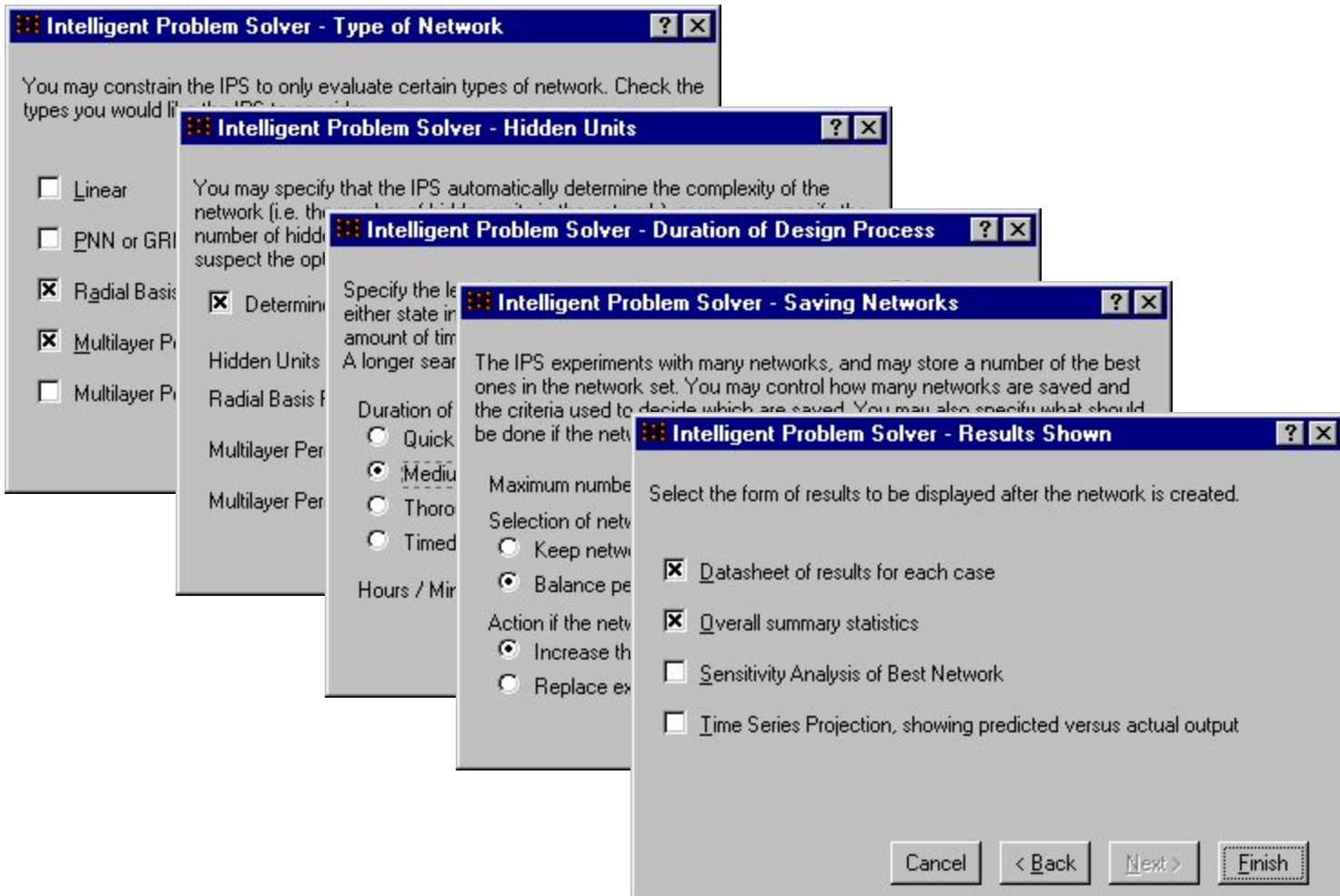
Variable	Value
ПРОДАЖИ	
Янв-1979	5020
Фев-1979	4873
Мар-1979	5460
Апр-1979	5590
Май-1979	6055
Июн-1979	6282
Июл-1979	6366
Авг-1979	6834
Сен-1979	6531
Окт-1979	6822
Ноя-1979	6777
Дек-1979	6905

**Выходная переменная
автоматически подается
на вход сети**

Intelligent problem solver



Intelligent problem solver





Data Set Editor (Retail)

Variables 1 1 Cases 133 120 0

ПРОДАЖИ	
Янв-1979	5026
Фев-1979	4873
Мар-1979	5460
Апр-1979	5590
Май-1979	6055
Июн-1979	6282
Июл-1979	6366
Авг-1979	6834
Сен-1979	6531
Окт-1979	6822
Ноя-1979	6777
Дек-1979	6905

Процесс поиска сети

Intelligent Problem Solver Messages

Intelligent Problem Solver, Time Series problem, 12 input steps
(select Options/Detailed Feedback for more detailed, technical feedback)
Input variables fixed
Network complexity will be determined automatically
The best 10 networks found (taking account of diversity) will be retained
Starting the search now...
0:00:21 An improved network has been found
0:00:51 An improved network has been found

Intelligent Problem Solver (Press ESC to Cancel the search)

0:01:09



Run Data Set

Outputs shown: **Variables** [Run] [Data Set]

RMS Error Train: 218.4 Verify: 461.9 Test: 0

	ПРОДАЖИ	Т. ПРОДАЖ	Е. ПРОДАЖ	Error
Авр-1995	14305.59	14186	119.5856	0.009371
Сен-1995	13118.64	13213	-94.35919	0.007394
Окт-1995	13801.3	13190	611.3029	0.0479
Ноя-1995	12741.05	12650	91.04846	0.0071349
Дек-1995	12840.51	12931	-90.48922	0.007091
Янв-1996	12834.57	12456	378.5716	0.02967

Regression Statistics

Variable: 1 [Run]

	Tr. ПРОДА	Ve. ПРОДА
Data Mean	10866.89	11608.5
Data S.D.	2580.426	3060.014
Error Mean	-2.540846	14.17079
Error S.D.	320.6128	463.5951
Abs E. Mean	251.4511	342.2935
S.D. Ratio	0.124248	0.151501
Correlation	0.9922536	0.9885846

Network Set Editor (Retail)

Current network: TC [Detail shown: Basic] [Options...]

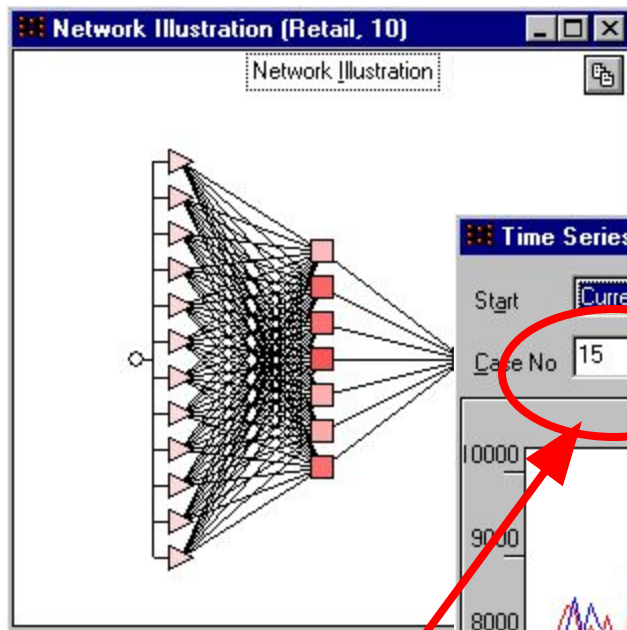
	Error	Inputs	Hidden	Performance
04	717.0334	1	7	0.2344311
05	654.8479	1	11	0.212822
06	543.6919	1	7	0.1783217
07	543.2743	1	5	0.1782253
08	525.7663	1	5	0.1725092
09	517.8769	1	9	0.169764
10*	461.8769	1	7	0.151501

PRODAJ Error (Train)

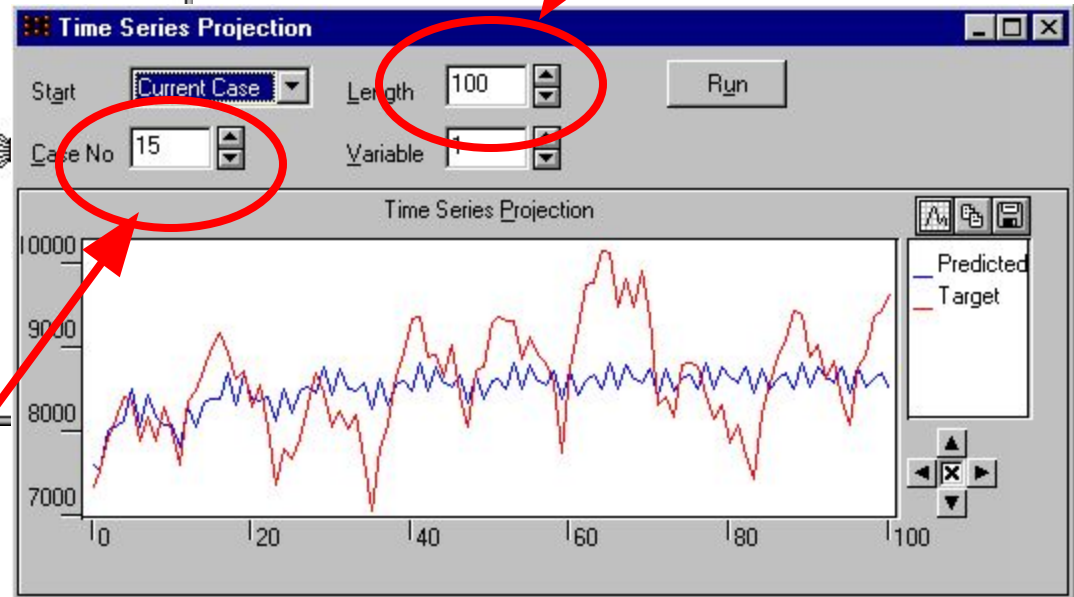
Variables: 1 1 Cases: 133 120 0

	ПРОДАЖИ
январь-1979	5026
февраль-1979	4873
март-1979	5460
апрель-1979	5590
май-1979	6055
июнь-1979	6282
июль-1979	6366
август-1979	6834

Обсуждение результатов



Глубина прогноза



Отправная точка

Train Statistics Run Options Window Help

Multilayer Perceptrons

- Back Propagation...
- Conjugate Gradients...
- Quasi-Newton...
- Levenberg-Marquardt...**
- Quick Propagation...
- Delta-Bar-Delta...

Radial Basis Functions...
 Linear
 Kohonen...
 Class labels...
 LVQ...
 Probabilistic...
 Generalized Regression...
 Principal Components...
 Auxiliary

Data
 Variables
 Янв-1979
 Фев-1979
 Мар-1979
 Апр-1979 5590
 Май-1979 6055

**Необходимо
переобучить
сеть**

Network Set Editor (Retail)

Current network: 10 Detail shown: Basic Options...

Type	Error	Inputs	Hidden
04	RBF 717.0334	1	7
05	RBF 654.8479	1	11
06	MLP 543.6919	1	7
07	MLP 543.2743	1	5
08	MLP 525.7663	1	5
09	MLP 517.3763	1	9
10 *	MLP 461.8769	1	7

Network Illustration (Retail, 10)

Network Illustration

Levenberg-Marquardt Training

Epochs: 1000 Train
 Cross verification Reinitialize
 Log Weights
 Stop
 Close

0:00:00



Data Set Editor (Retail)

Variables: 1 | 1 | Cases: 133 | 120 | 0

ПРОДАЖИ	
Янв-1979	5026
Фев-1979	4873
Мар-1979	5460
Апр-1979	5590
Май-1979	6055

Levenberg-Marquardt Training

Epochs: 1000

Cross verification

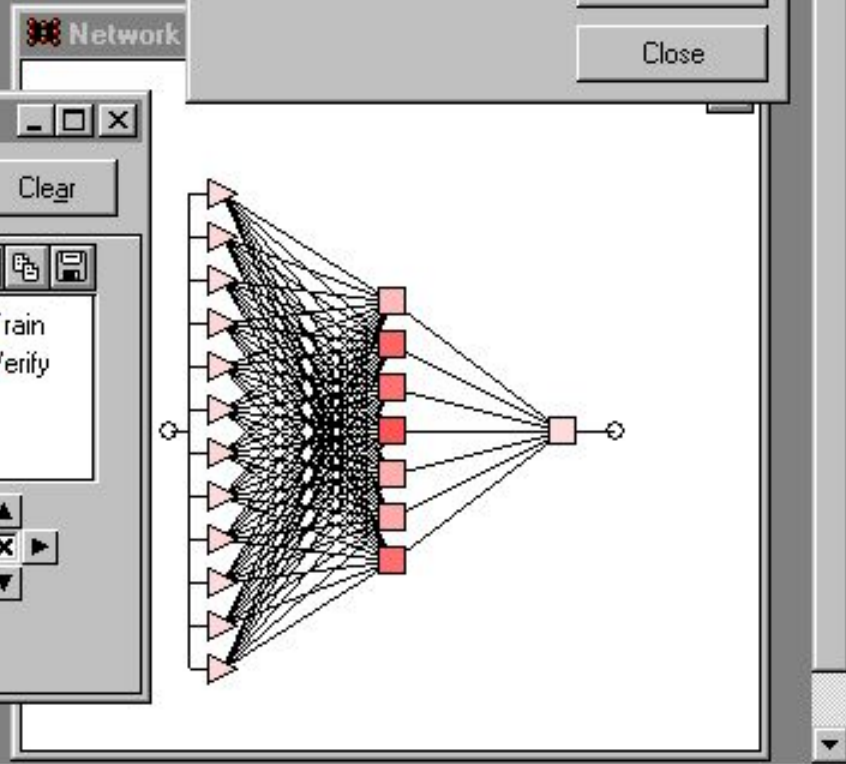
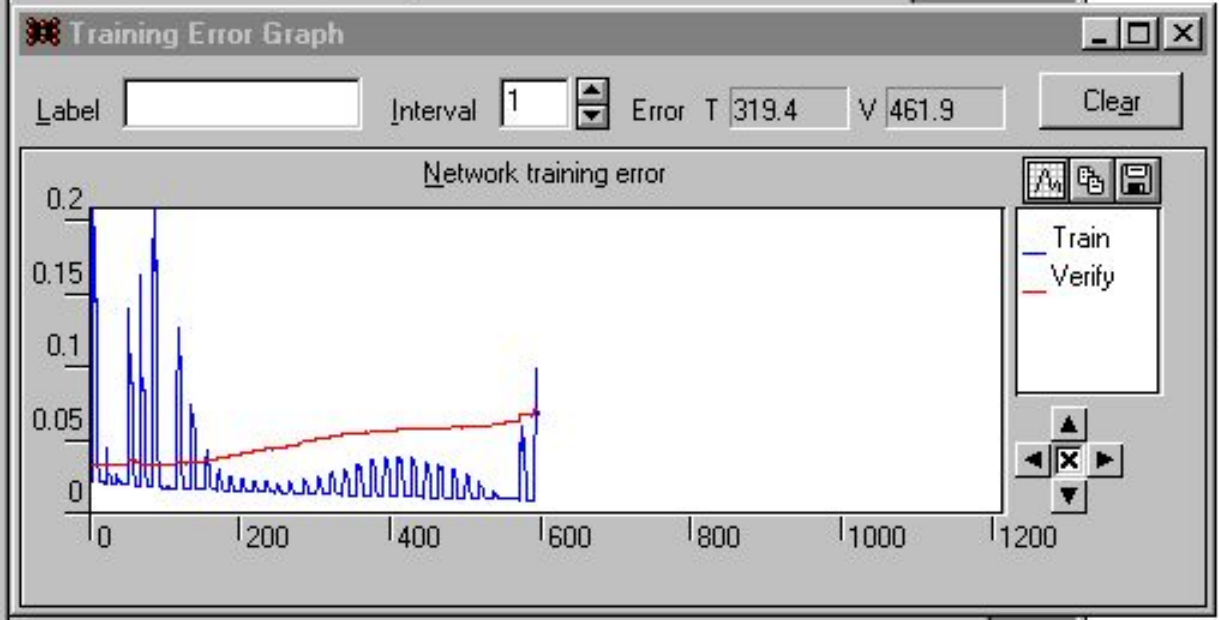
Train

Reinitialize

Log Weights

Stop

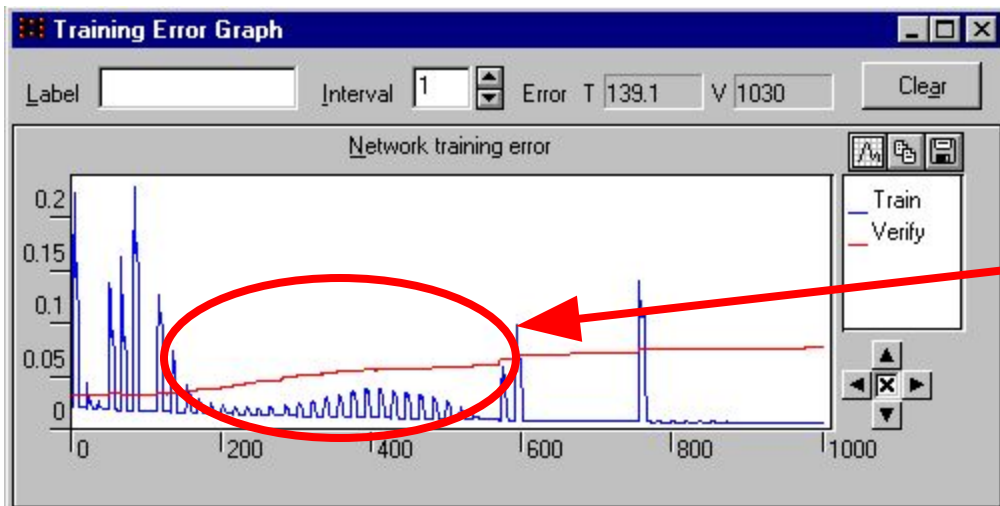
Close



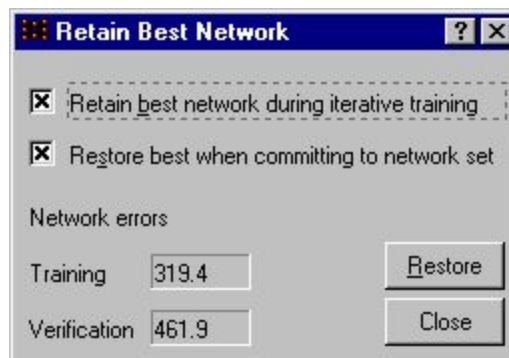
Starting Levenberg-Marquardt training

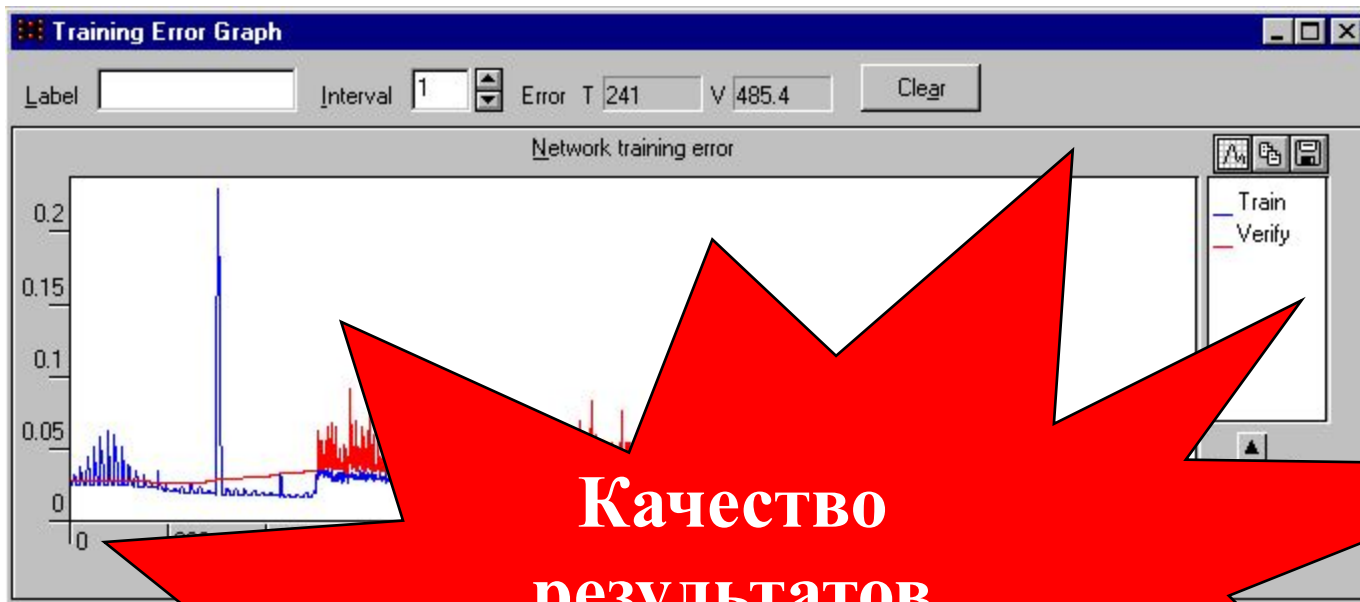
0:01:19

Результаты



**Возможно
переобучение!**





**Качество
результатов
сравнимо
с классическими
методами.**





Качество прогноза

Для повышения качества прогноза рекомендуется добавить к исходной переменной ряд, определяемый как

$$Dy(t) = y(t) - y(t-1)$$

**Точность прогноза
увеличилась на порядок!**



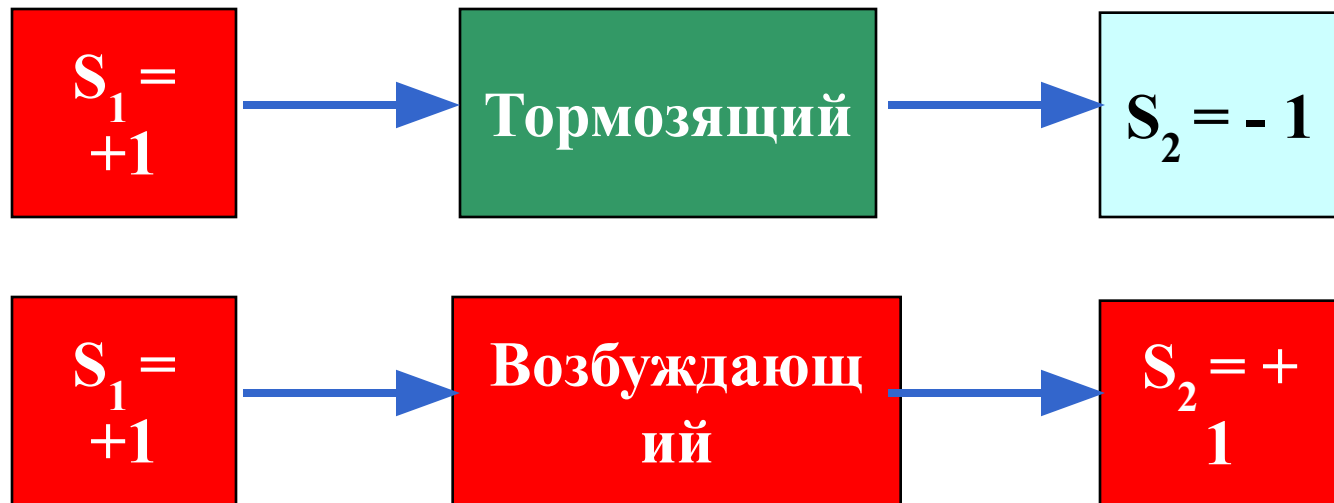
Модель распределенной нейронной памяти

Каждый нейрон может находиться в двух состояниях:

- $S_1 = +1$ - возбужденное
- $S_2 = -1$ - покой

Нейроны связаны между собой синаптическими связями, которые бывают возбуждающие и тормозящие.

Модель распределенной нейронной памяти



Связь можно описывать коэффициентом: $J_{ij} = \pm 1$


$$E_{12} = J_{12}S_1S_2$$

- потенциальная энергия связи

Любая система предоставленная самой себе
стремится к минимуму своей потенциальной
энергии.

$$\sum_{ij} E_{ij} \rightarrow \min$$



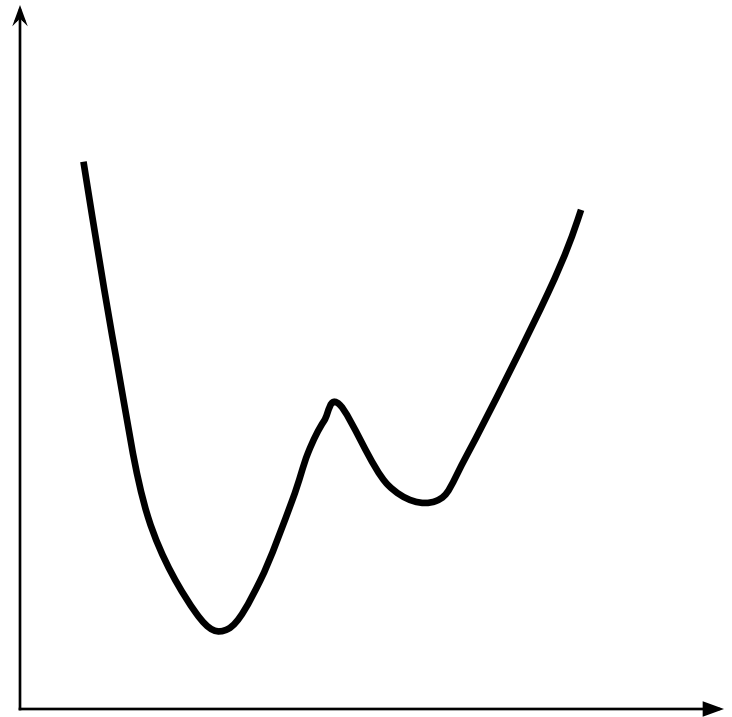
Образ

$$J_{ij} = \frac{1}{N} \sum_{\alpha} \zeta_i^{\alpha} \zeta_j^{\alpha}, i = j, J_{ii} = 0, J_{ij} = J_{ji}$$

**Модель Хопфилда коэффициентов
межнейронных связей, когда в сети
запомнено
р образов.**

**Каждому образу соответствует
локальный энергетический минимум!**

Бимодальный образ



Качественный вид потенциальной функции Хопфилда



Модель Хакена

**Образы описываются
параметрами порядка: d_1 и d_2**

**Переменные, описывающие степень
насыщения внимания: k_1 и k_2**

Уравнения насыщения

$$d'_1 = d_1(k_1 - Ad_1^2 - Bd_2^2)$$

$$d'_2 = d_2(k_2 - Bd_1^2 - Ad_2^2)$$

$$k'_2 = g(1 - k_2 - d_2^2)$$

$$k'_1 = g(1 - k_1 - d_1^2)$$

**При некотором соотношении между
константами**

A, B, g имеет место осцилляция внимания!

Период колебаний

- При зрительном восприятии:

$$T = 10\text{с}$$

- При смысловой неоднозначности:
«Продается собака. Неприхотлива в еде. Любит детей.»

$$T = 0.1\text{с}$$

Разница объясняется
существенной разницей нервного
вещества, вовлеченного в эти
процессы.



Мы обсудили

- Проблему неоднозначности в искусстве
- Биологические нейронные сети
- Математические модели нейронных сетей
- Возможности пакета SNN
- Задачу: прогноз результатов выборов президента США



Мы обсудили

- Задачу: выявление показателей, влияющих на валовую прибыль предприятия (регрессионная модель)
- Задачу: прогнозирование временного ряда
- Способ оценки периода колебаний зрительных образов в сознании

